

# Identification of Musical Instruments Using MFCC Features



Sushen R. Gulhane, D. Shirbahadurkar Suresh, and S. Badhe Sanjay

## 1 Introduction

We have worked on the most popular area of research i.e. musical instrument identification. In recent days most of the musical data is available in the digital form. In such cases, automatic musical instrument identification becomes more important and popular. Monophonic, polyphonic, homophonic are the different types of music signals [1]. Where the monophonic type of music signal means only one musical instrument is playing at that time and the audio sound signal is of only one instrument playing a single melody. Our work deals with the identification of musical instrument wherein the audio sound file is of the monophonic type which is the texture of the music. Similar musical instruments are having almost similar kind of features. A group of sound samples is taken. Features of these group are extracted using different feature extraction techniques. There are many feature extraction techniques to extract features of audio sound signals [2]. In our work, we are using MFCC feature extraction method. Some algorithms are used to compare various features from audio samples. These algorithms are called as classifiers. In our work, K-Nearest Neighbor (KNN) and Support Vector Machine (SVM), these two classifiers have been used to identify the musical instrument.

Our objectives for the research are to recognize musical instrument from an audio signal and to check which classifier can give better identification. In this whole work of identification, there are three tasks. First one is pre-processing of the input music signal, the second one is feature extraction from sound signals and the last one is recognition of instrument [3].

---

Sushen R. Gulhane (✉) · S. Badhe Sanjay  
DYPIT, DYPCOE (SPPU), Pune, India

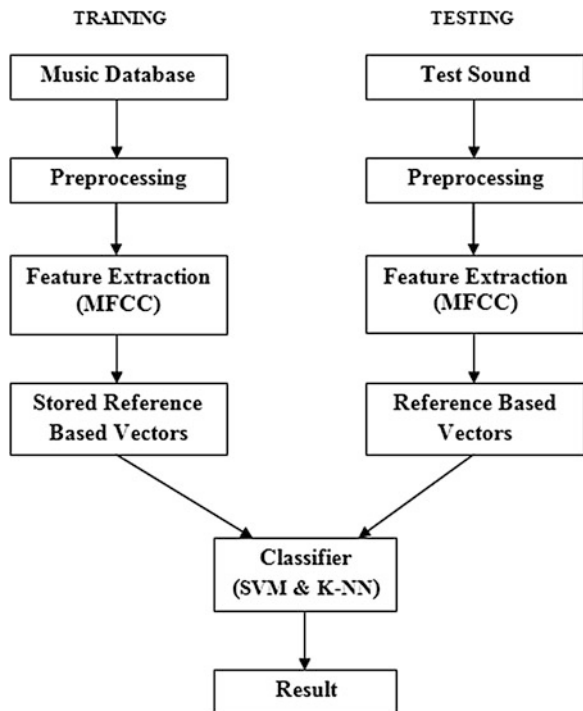
D. Shirbahadurkar Suresh  
DYPIT, Zeel COE (SPPU), Pune, India

## 2 System Overview

Our proposed system of musical instrument identification divides and flow in two phases, training phase flow, and the testing phase flow as shown in Fig. 1. In the training phase, different audio samples of musical instruments are fed as input to the musical instrument identification system. MFCC features are extracted from these audio samples by using MFCC feature extraction technique. Those features are stored in an array-matrix format in excel which is called as features array. These extracted features are fed to the system for training purpose using specified classifier for further classification process [4]. Similarly, in the testing phase flow, an unknown audio sound signal file is given as an input to the system. Then the MFCC features are extracted of the same audio sound signal file and results are compared with the reference results obtained in the training phase flow and the new signal is then classified by using the same classifier [5]. The purpose of the proposed work is to find out better classifier and combinations of feature extraction methods with classifiers. In the first step, the musical instruments sound database is created. Finally, the input sample is compared with reference samples to find out the type of instrument.

There are different machine learning techniques that have been used for musical instrument identification [6]. Two main types of learning techniques are Supervised

**Fig. 1** General block diagram of proposed musical instrument recognition system



and Unsupervised. In a supervised learning technique, there is a set of audio sound signals for the purpose of reference called the training vector. In this method, final resultant outputs are available and we have to get the output almost the same as a target output. In unsupervised learning, outputs are finding out with the help of inputs only. There are no target outputs available. In our system, we are using a supervised learning technique.

Preprocessing is the first step in instrument identification which is used to increase the efficiency of overall system performance [7]. In a preprocessing step, it contains two steps of reading sounds and noise removal. Noise removal is used to increase the strength of the signal and also to increase the accuracy of recognition [8]. Next part of the project is feature extraction and then classification. There are a lot of feature extraction techniques. There are different types of features for musical instrument identification. A study is to be carried out for the specific features selection and then it will be used. In this paper, the MFCC feature extraction method is used. For the identification of musical instruments, the need is to extract meaningful information. The audio signal will be consisting of a lot of information where we will not be needing all the information for the identification. So by using feature extraction methods, meaningful information is to be extracted and will be used for further processing [9]. Most of the time, it happens that audio data to be provided to the classifier is too large with less information, which reduces the efficiency of the classifier, in that case, meaningful data from the audio data samples are to be extracted and used. Mostly that will be the collection of no. Of features hence called a feature vector [10, 11]. These sets of information related to features are then used to identify the musical instrument. Further, in the third step, the recognition process is done by using classifiers.

### 3 System Design Methodology

Here we are working on musical instrument identification where monophonic audio samples have been considered along with the MFCC feature extraction method, two different classifiers, and the database system. We have given the details of the different techniques used in this project work.

#### 3.1 *Mel-Frequency Cepstral Coefficients (MFCC)*

There are a lot of features to be extracted from the audio sample out of which MFCC features give the best result for the identification. MFCCs always represent spectral coefficients in a sound. Each coefficient is having a specific value for each part of the sound signal. Now a day MFCC is widely used in instrument and speaker identification systems. Figure 2 presents the block diagram for the detail process of extracting MFCC features.

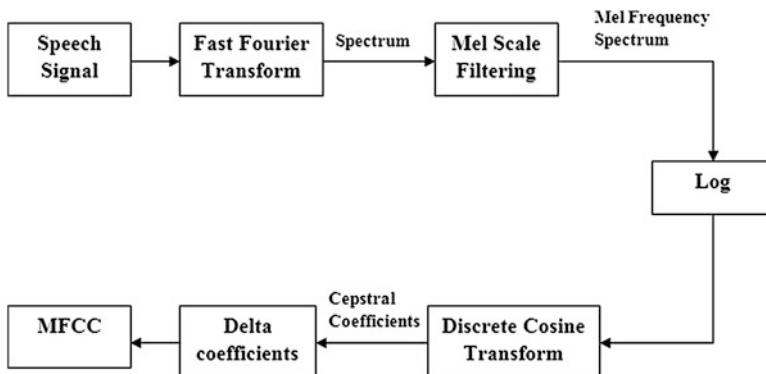


Fig. 2 Block diagram for extracting MFCC features

### 3.1.1 Database

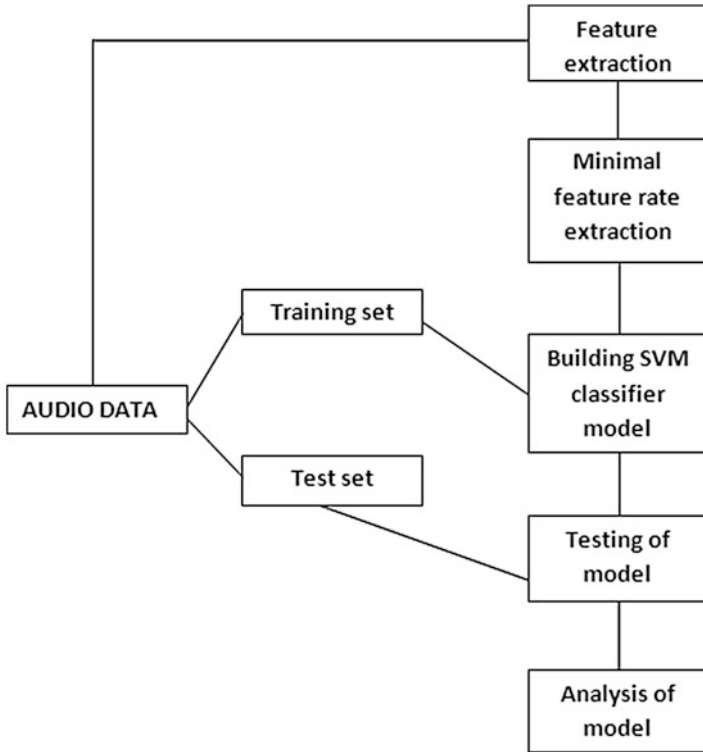
The database is a well-maintained collection of information which is organized in a particular way so that a program can quickly and easily select any required part of data [12]. In our work, we have to make a database of audio samples for different musical instruments like piano, sitar, guitar, etc. In this work, we are using an in-house database so that we can record any number of data for each instrument. Audio samples of different instruments are recorded in .wav format with the same size. From the recording sample database, 70% database is used for training purpose and 30% database is used for testing purpose.

### 3.1.2 Support Vector Machine (SVM)

There are no. of classifiers to classify the data out of which, SVM is a classifier which comes under supervised learning techniques. SVM is used in various applications like face, character, handwriting recognition. To transform the data it uses a technique which is called the kernel trick. SVM finds the minimum boundary between the possible outputs using those data transformations. Though SVM design is complex but it gives the best results as compared to other classifiers. SVM classifier algorithm process is as shown by Fig. 3.

### 3.1.3 K-Nearest Neighbor Algorithm

In K-nearest neighbor algorithm it finds objects of the same nature and makes a grouping of them. In K-NN algorithm process objects of the same category or having the same features should be closer in distance. It is a type of instance-based learning technique and it gives the class of a new test data based on the data samples



**Fig. 3** Audio classification model using SVM

in the database. Out of various machine learning algorithm, KNN is the easiest machine learning algorithm. The flowchart for the K-NN algorithm is as shown in Fig. 4.

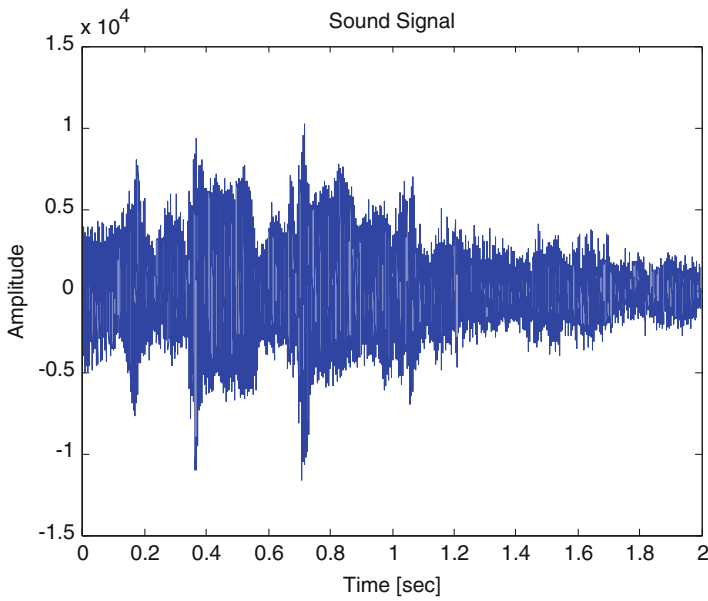
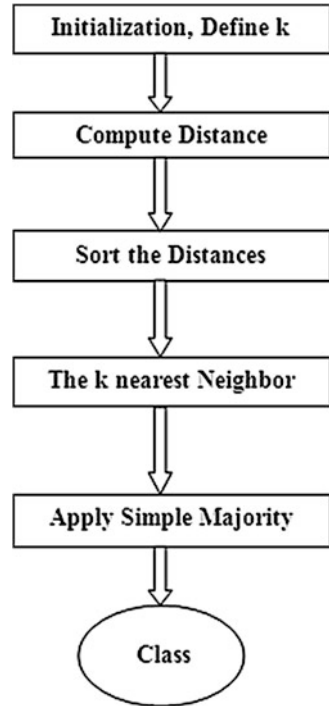
### 4 Results

We have used MFCC for feature extraction and following figures showing some results on input sound signal in the process of converting it into Mel frequency spectrum. Figure 5 shows the input signal waveform which is in the form of sound.

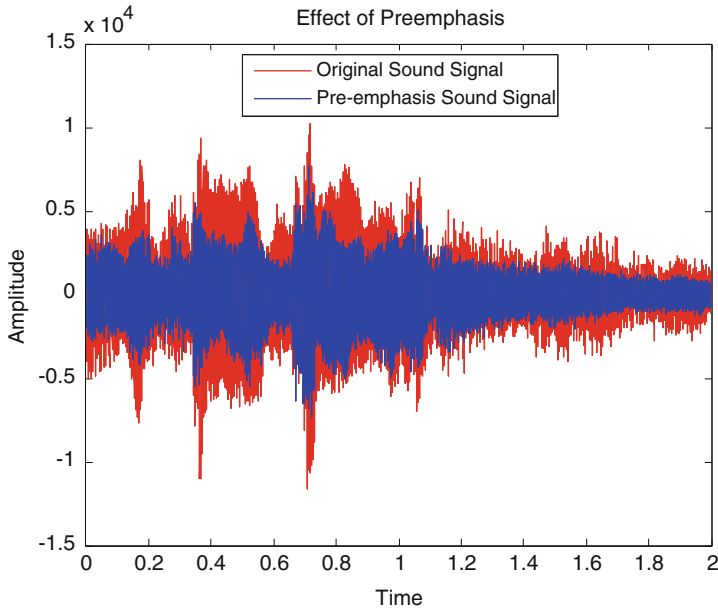
Figure 6 shows the effect of pre-emphasis that is nothing but noise removal on input sound signal.  $S(n)$  represent the input sound signal,  $n$  be the time coefficient, and  $\alpha = 0.97$ .

$$\hat{S}(n) = S(n) - \alpha.S(n - 1) \tag{1}$$

**Fig. 4** Flowchart for K-NN classifier algorithm



**Fig. 5** Input sound signal



**Fig. 6** Pre-emphasis filtering

Figure 7 shows hamming window with 64 samples. A Hamming window is  $A = \pi r^2$  generally used to prevent dramatic changes in a window, as follows:

$$W(n) = \begin{cases} 0.54 - 0.46 \times \cos\left(\frac{2n\pi}{N-1}\right), & 0 \leq n \leq N - 1 \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

To remove additional high-frequency signals on both sides of frames of sound signals window processing is used. It highlights major signals at the center of the frame. To better observe sound signal characteristics, we plot different slots in a signal like frames which are generally with 50%. It is referred to as framing as shown in Fig. 8.

Figure 9 shows tri-filter bank signal. It reduces the frequency scale. Triangular filter bank consists of a no. of triangular bandpass filters.

Discrete cosine transform is used to transform frequency to the time-domain signal. It results in Mel Frequency Cepstrum as shown in Fig. 10 by using log filter bank energies as shown in Fig. 11. In the identification system of sound, the most commonly used simple frequency scale transformation equation is Mel frequency which is shown below.

$$\text{Mel} = 2595 \times \log\left(1 + \frac{f}{700}\right) \quad (3)$$

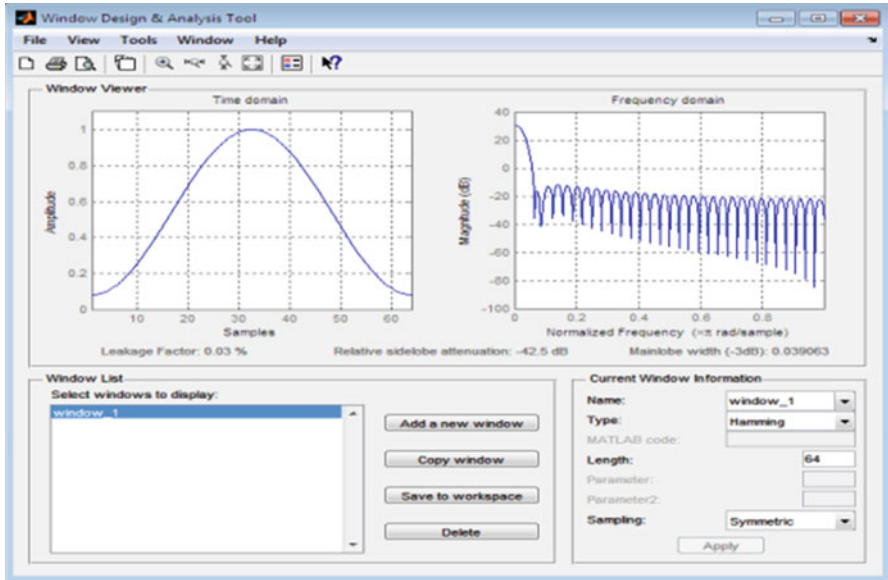


Fig. 7 Hamming window

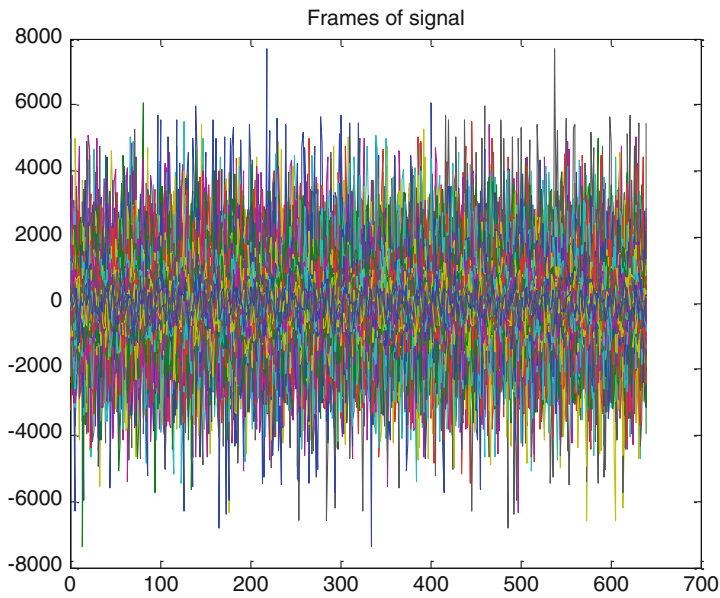


Fig. 8 Different frames of signal



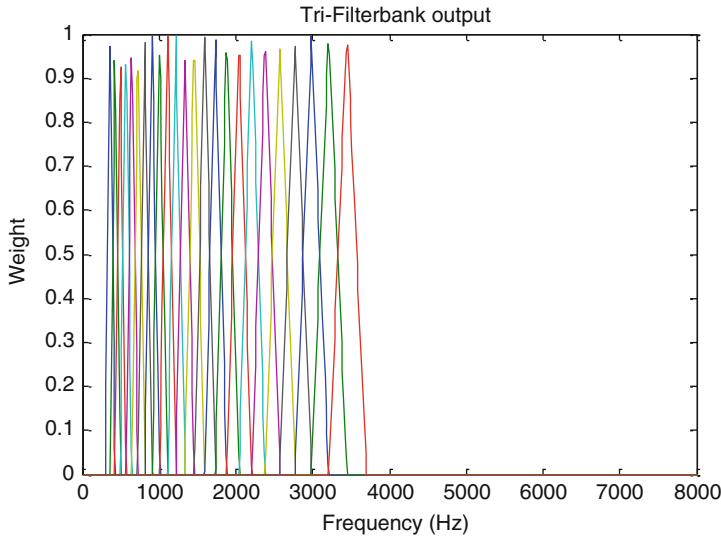
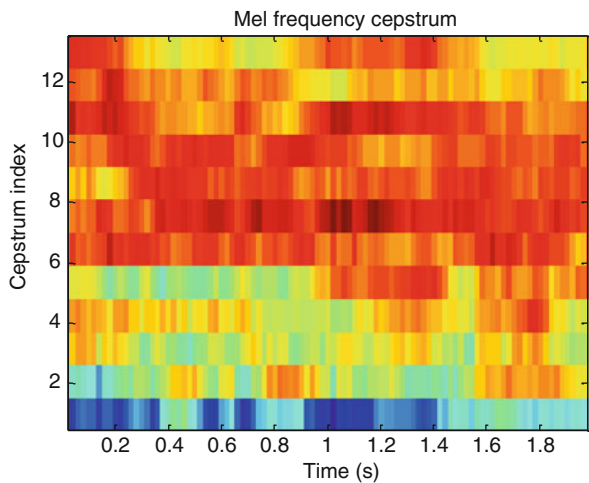


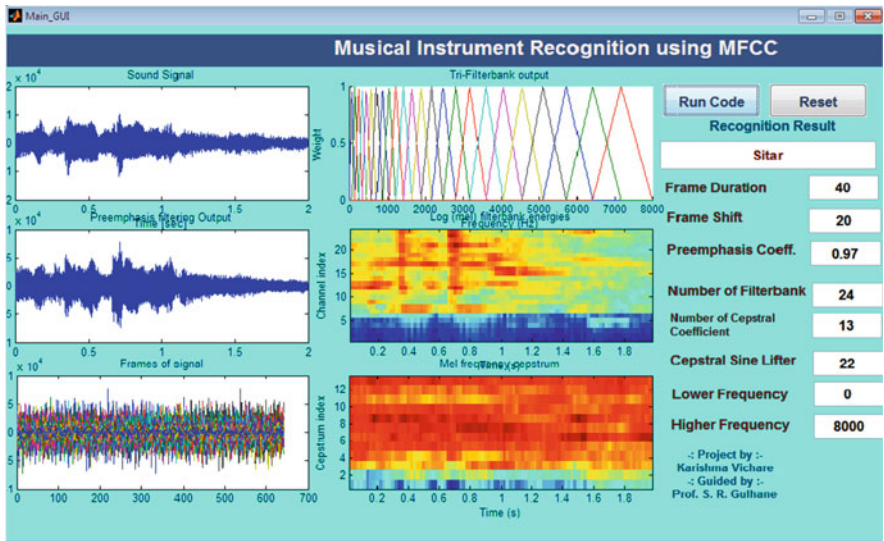
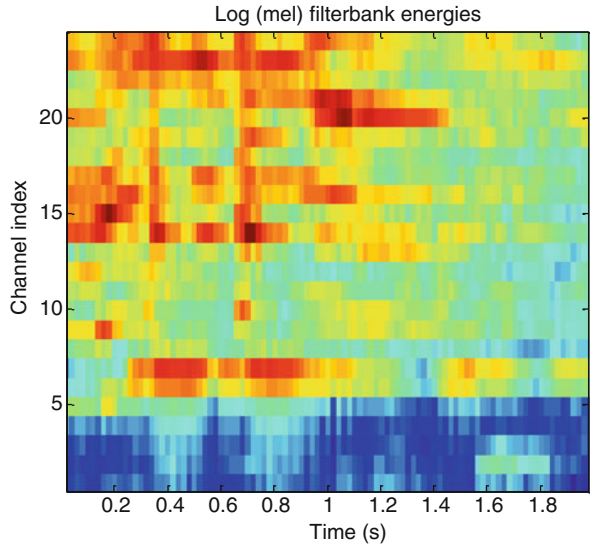
Fig. 9 Tri-filter bank signal

Fig. 10 Mel frequency cepstrum



We have made a Graphical User Interface for our recognition system which gives different results related to MFCC in the form of waveforms as shown in Fig. 12. It contains all related parameters which we have used in the process of feature extraction and classification. By using Run code button we can recognize the instrument played in the selected sound file. Reset button clears all previous output waveforms. The final recognition result which is the name of the instrument has displayed by the GUI after the testing phase.

**Fig. 11** Log filter bank energies



**Fig. 12** GUI for musical instrument recognition

K-NN algorithm gives various results for various values of K and the number of instruments. Tables 1, 2, and 3 gives confusion matrices for the K-NN algorithm. Table 1 gives various accuracy percentage values for different instruments with the value of K is 5. Table 1 gives less accuracy as compared to Tables 2 and 3. Table 2 gives various accuracy percentage values for different instruments with the value of K is 3. Table 2 gives less accuracy as compared to Table 3 and more than Table 1.

**Table 1** Confusion Matrix for K-NN algorithm with K = 5

For K = 5	Drum	Sitar	Flute	Piano	Tabla	Accuracy
Drum	27	0	0	4	4	77.14
Sitar	1	20	0	4	0	80
Flute	1	3	18	3	0	72
Piano	0	1	0	24	0	96
Tabla	1	0	0	1	23	92

**Table 2** Confusion Matrix for K-NN algorithm with K = 3

For K = 3	Drum	Sitar	Flute	Piano	Tabla	Accuracy
Drum	31	0	0	2	2	88.5714
Sitar	1	23	0	1	0	92
Flute	1	2	21	1	0	84
Piano	0	1	0	24	0	96
Tabla	0	1	0	0	24	96

**Table 3** Confusion Matrix for K-NN algorithm with K = 3

For K = 1	Drum	Sitar	Flute	Piano	Tabla	Accuracy
Drum	35	0	0	0	0	100
Sitar	0	25	0	0	0	100
Flute	0	0	25	0	0	100
Piano	0	0	0	25	0	100
Tabla	0	0	0	0	25	100

Table 3 gives various accuracy percentage values for different instruments with the value of K is 1. Table 3 gives maximum accuracy as compared to Table 1 and 2. It gives 100% accuracy for various numbers of instruments.

## 5 Conclusion

To implement Musical Instrument Identification System, we have performed a lot of experimentation considering MFCC feature extraction techniques and SVM & KNN as classifiers and concluded that MFCC gives better performance with SVM classifier as compared to K-NN classifiers.

The system gives better accuracy in identification of musical instrument and status finding. The future work of this system is to use more audio features, such as PLP or Wavelet Transform and to make instrument recognition for polyphonic music.

## References

1. Athanasia Zlatintsi and Petros Maragos “Multiscale Fractal Analysis of Musical Instrument Signals With Application to Recognition”, IEEE Transactions On Audio, Speech, And Language Processing, Vol. 21, No. 4, April 2013.
2. S. R. Sankaye and U. S. Tandon, “Indian Musical Instrument Recognition based MFCC Feature Set”, IOSR Journal of Computer Engineering (IOSR-JCE).
3. Naktode Dipali Ravi and Daulappa G. Bhalke “Musical Instrument Information Retrieval Using Neural Network”, 2016 IEEE International Conference on Advances in Electronics, Communication and Computer Technology (ICAECCT).
4. Miss. Karishma Vichare 1, Prof. S. R. Gulhane 2 “MUSICAL INSTRUMENT IDENTIFICATION AND STATUS FINDING WITH MFCC”, International Journal of Advance Engineering and Research Development, Volume 5, Issue 04, April-2018.
5. Priyanka S. Jadhav, “Classification of Musical Instruments sounds by Using MFCC and Timbral Audio Descriptors”, International Journal on Recent and Innovation Trends in Computing and Communication, July 2015 Volume 3 Iss. 7.
6. S. R. Sankaye and U. S. Tandon, “Human Perception and Recognition of Musical Instruments”.
7. Lindasalwa Muda, Mumtaj Begam and I. Elamvazuthi, “Voice Recognition Algorithms using Mel Frequency Cepstral Coefficient (MFCC) and Dynamic Time Warping (DTW) Techniques”, Journal Of Computing, March 2010, ISSN 2151-9617
8. Babak Toghiani-Rizi and Marcus Windmark “Musical Instrument Recognition Using Their Distinctive Characteristics in Artificial Neural Networks”, May 2017.
9. A. S. Patankar, “Musical Instrument Recognition Techniques, IJRBAT, Jan. 2017.
10. Antti Eronen and Anssi Klapuri, “Musical Instrument Recognition Using Cepstral Coefficients and Temporal Features”, P.O. Box 553, FIN-33101 Tampere, Finland.
11. Monica S. Nagawade, Varsha R. Ratnaparkhe, “Musical Instrument Identification using MFCC”, 2017 2nd IEEE International Conference On Recent Trends in Electronics Information & Communication Technology (RTEICT), May 19–20, 2017, India.
12. Jeremiah D. Deng, *Member, IEEE*, Christian Simmermacher, and Stephen Cranefield, “A Study on Feature Analysis for Musical Instrument Classification”, IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS—PART B: CYBERNETICS, VOL. 38, NO. 2, APRIL 2008.